AP680-1

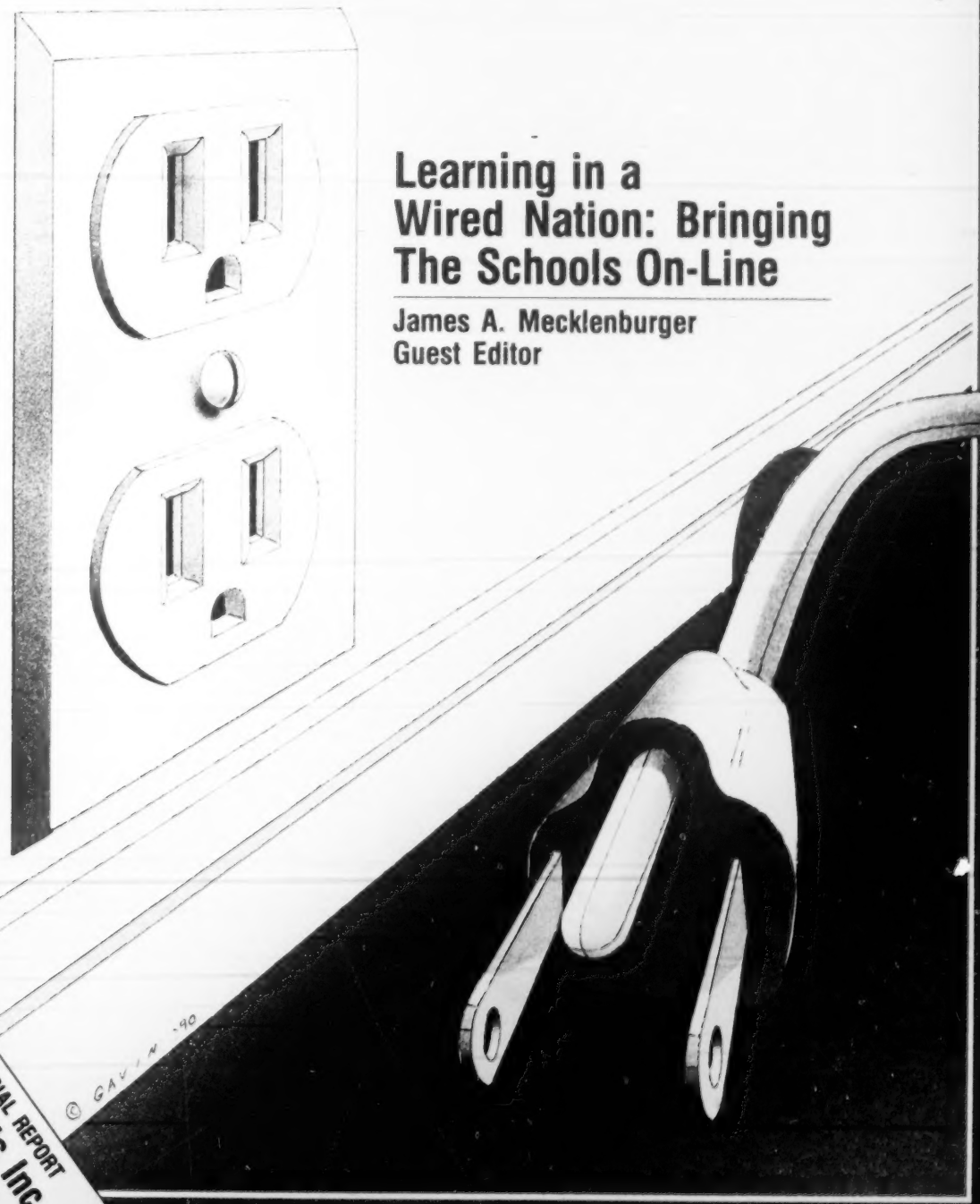
PHI DELTA

OCTOBER 1990

KAPPAN

Learning in a Wired Nation: Bringing The Schools On-Line

James A. Mecklenburger
Guest Editor



A KAPPAN SPECIAL REPORT

College Sports Inc.

Murray Sperber

Phi Delta October 1990 • Volume 72 • Number 2

KAPPAN

FEATURES

104 **Educational Technology Is Not Enough,** *by James A. Mecklenburger*

109 **The Wired Classroom: American Education Goes On-Line,** *by Bruce Watson*

113 **The New Generation of Information Systems,** *by Peter Grunwald*

115 **Is There a Role for Educators in Telecommunications Policy?,** *by Shelly Weinstein and Susanne A. Roschwalb*

118 **Implementing an Integrated Instructional System: Critical Issues,** *by Mark Sherry*

121 **Will Technology Transform Education or Will the Schools Co-opt Technology?,** *by Eliot Levinson*

127 **The Nation's Report Card Goes Home: Good News and Bad About Trends in Achievement,** *by Robert L. Linn and Stephen B. Dunbar*

134 **Continuing to Cooperate: Prerequisites for Persistence,** *by Elizabeth G. Cohen*

KAPPAN Special Report – College Sports Inc.: The Athletic Department vs. the University, *by Murray Sperber*

139 **The Evolving Effort to Improve Schools: Pseudo-Reform, Incremental Reform, and Restructuring,** *by Mary Anne Raywid*

144 **Stepfamilies: Second Chance or Second-Rate?,** *by Jo Barney*

148 **Does It Make a Difference If You Change the Structure? School-Within-a-School,** *by Melinda Nixon Nickle, Fran Carter Flynt, Stephen Douglass Poynter, and James A. Rees, Jr.*

153 **How to Keep Even Your Best Friend from Getting Tenure,** *by Alleen Pace Nilsen and Sandra Luehrsen*

156 **An Educational Safari,** *by Jamieson A. McKenzie*

DEPARTMENTS

99 **THE EDITOR'S PAGE,** Afraid of the Dark?

100 **WASHINGTON COMMENTARY,** *by Anne C. Lewis*

102 **STATELINE,** *by Chris Pipho*

158 **PAR,** Cooperative Learning: Fad or Foundation for Learning?, *by Deborah Burnett Strother*

164 **DE JURE,** Handicapped Parents, *by Perry A. Zirkel*

168 **IN CANADA,** French Education Rights Upheld by Supreme Court, *by Tom McConaghy*

170 **PROTOTYPES,** BioPrep: An Academic Honors Program for Rural Students, *by Larry Rainey*

172 **BACKTALK**

Cover Illustration: Kerry Gavin

Afraid of the Dark?

FOR THOSE wishing to take part in the conversation on school reform in recent years, the password has been *restructuring*. And the *Kappan* has dealt with that topic repeatedly. We've talked about tampering with course schedules and time constraints, about blurring (or eradicating) the boundaries between disciplines, about empowering teachers and principals, about moving beyond the traditional grade-level structure. But the changes that we've suggested have all been on the level of rearranging furniture.

The special section on technology in this month's *Kappan*, by contrast, pushes at the very walls of the school. As guest editor James Mecklenburger points out, it's time to stop thinking about "technology in education" and to start thinking about education within the context of the new technologies.

We're talking *radical* restructuring here – systemic change. And the central question becomes: Who will best use technology in the service of education, professional educators or the private sector? As the articles in this month's issue point out, private providers are currently in the lead – and the education establishment is still absent from the debates about telecommunications policy.

Cynics might say this month's special section is for dreamers. And clearly, there are plenty of barriers to keep professional educators from tapping the transformative powers of technology in any large-scale way.

The major hurdle is the high cost of the necessary hardware and software. In a spring 1990 special issue, *Fortune* pointed out that American schools have only one computer for every 30 students. To reduce that ratio to one for every three students, according to the Office of Technology Assessment, the U.S. would have to spend an extra $4 billion annually for many years.

Then there's the technophobia that afflicts a vast majority of American teachers. Expensive staff development efforts and major changes in preservice teacher education would be required to overcome this barrier.

And that's not even to mention the issue of equity, in an era still plagued by immense financial disparities among school districts. Clearly, bake sales and car washes will not suffice to finance this restructuring effort.

The financial ramifications are sobering. But they are no more sobering than Eliot Levinson's warning in this issue about the difficulty of implementing radical organizational and structural innovation of any kind in our education system.

Just suppose for a moment, though, that the schools could surmount these hurdles. We'd have active, learner-centered, lifelong, interdisciplinary learning. We'd have genuine equality of educational opportunity. We'd have real individualization of instruction. We'd have ways of dealing effectively with adult illiterates and with dropouts in search of a second chance. We'd have new and powerful solutions to a whole raft of problems that now bedevil us.

We live in a wired nation, as Shelly Weinstein and Susanne Roschwalb point out in this issue, so the technology is at hand. But our schools are still unplugged. It's time to take steps to make sure that the dark doesn't catch us. – PBG

KAPPAN STAFF

PAULINE B. GOUGH, Editor
BRUCE M. SMITH, Managing Editor
RISÉ WILLIAMSON, Assistant Editor
STANLEY M. ELAM, Contributing Editor
CAROL BUCHERI, Design/Production Director, Advertising/Circulation Manager
VICTORIA VOELKER, Designer
TERRI HAMPTON, Permissions
SHEILA WAY and DEBBIE WEBB, Compositors

PHI DELTA KAPPAN USPS 429-840

(ISSN 0031-7217) Published monthly, except July and August, by Phi Delta Kappa, Inc., Eighth & Union, P.O. Box 789, Bloomington, IN 47402. Subscription rate, $30 per year, domestic; $32.50 per year, foreign. Single copies $3.50 each; remit with order. Indexed in *Education Index* and in *Current Index to Journals in Education*, available on microfilm, University Microfilms, Inc., Ann Arbor, Mich. Second-class postage paid at Bloomington, Ind., and at additional mailing offices. *Postmaster*: Send address changes to *Phi Delta Kappan*, P.O. Box 789, Bloomington, IN 47402. Copyright 1990 by Phi Delta Kappa, Inc.

HEADQUARTERS STAFF

LOWELL C. ROSE, Executive Director; DON L. PARK, Associate Executive Director Program; LARRY W. BARBER, Director, Center on Evaluation, Development, and Research (CEDR); DEREK L. BURLESON, Editor of Special Publications; HOWARD D. HILL, Director of Chapter Programs; ROBERT E. McDANIEL, Director of Administrative Services; NEVILLE L. ROBERTSON, Director, Center for Dissemination of Innovative Programs; JACK FRYMIER, Senior Fellow; WILMER K. BUGHER, Director of Special Projects; VLADIMIR BEKTESH, Media Specialist; SHARI BRADLEY, Marketing and Training Coordinator; WILLARD DUCKETT, Assistant Director, CEDR; PHILLIP HARRIS, Director, Center for Professional Development; MONICA OVERMAN, Research Assistant, CEDR; DAVID M. RUETSCHLIN, Staff Associate, Special Publications; DEBORAH BURNETT STROTHER, CEDR Editor; SANDRA WEITH, Administrative Assistant, Administration; TERRY L. WIEDMER, Assistant for Membership/Public Information

BOARD OF DIRECTORS

CAROL O'CONNELL, President, Education Consultant, 7523 Elin Ct., Dayton, OH 45415
JACK KOSOY, President Elect, Admin., Adult Educ., L.A. Unified School Dist., 16539 Bircher St., Granada Hills, CA 91344
JOHN V. ROWLEY, Past President, Principal, Mt. Pleasant Middle School, Livingston, NJ 07039
ROGER W. ANDERSON, Vice President, Professor of Education, Ret., Luther College, Decorah, IA 52101
DOUGLAS BEDIENT, Vice President, Director, Learning Resources Service, Southern Illinois Univ. at Carbondale, Carbondale, IL 62901
SHIRLEY N. ROBARDS, Vice President, Director of Student Teaching, Univ. of Tulsa, Tulsa, OK 74104
DONALD H. DUNCAN, Director, Dist. I, Professor of Special Educ., Western Oregon State College, Monmouth, OR 97361
ROBERT L. MAULLER, Director, Dist. II, L.A. Unified School Dist. Admin., Ret., 550 Allendale Rd., Pasadena, CA 91106
J. ELDON YUNG, Director, Dist. III, Professor of Graphics, Central Missouri State Univ., Warrensburg, MO 64093
RICHARD L. KOLOWSKI, Director, Dist. IV, Director of Secondary Curriculum, Millard Public Schools, 5606 S. 147 St., Omaha, NE 68137
SONJA LOU CLARY, Director, Dist. V; Associate Dean, National-Louis University, 2840 N. Sheridan Rd., Chicago, IL 60201
JAMES V. FOGARTY, JR., Director, Dist. VI, Director, Special Educ., BOCES 2, Suffolk County, NY 11720
GEORGE M. THOMAS, Director, Dist. VII, Associate Professor, Mississippi State Univ., 5520 Highway 19 North, Meridian, MS 39307
GEORGE KERSEY, JR., Director, Dist. VIII, Principal, Carter-Lawrence Middle School, Nashville, TN 37203

EDITORIAL CONSULTANTS

TERREL H. BELL, Professor Emeritus of Educational Administration, Univ. of Utah, Salt Lake City
JOMILLS H. BRADDOCK II, Director, Center for Research on Effective Schooling for Disadvantaged Students, Johns Hopkins Univ., Baltimore
LARRY CUBAN, Professor and Assoc. Dean, School of Education, Stanford Univ., Stanford, Calif.
DENIS DOYLE, Senior Research Fellow, Hudson Institute, Alexandria, Va.
MARY HATWOOD FUTRELL, Senior Fellow and Assoc. Director of the George Washington Univ. Center for the Study of Education and National Development, Washington, D.C.
MARTIN HABERMAN, Professor, Univ. of Wisconsin, Milwaukee
ANN LIEBERMAN, Professor, Univ. of Washington, Seattle
JEROME T. MURPHY, Professor of Education and Assoc. Dean of the Graduate School of Education, Harvard Univ., Cambridge, Mass.
JOSEPH H. NATHAN, Senior Fellow, Hubert H. Humphrey Institute of Public Affairs, Univ. of Minnesota, Minneapolis
JOHN V. ROWLEY, Past President, PDK

Reprints of articles appearing in the *Kappan* may be obtained from the business office: 1-2 pp., $8; 3-4 pp., $13; 5-6 pp., $19; and 7-8 pp., $25 for 100 copies. On larger orders, write for price schedule. Please remit with order. Address all purchase orders and address changes to Director of Administrative Services. Allow six weeks for processing address changes. Address all editorial correspondence to *Phi Delta Kappan*, P.O. Box 789, Bloomington, IN 47402. Ph. 812/339-1156. Please include postpaid, self-addressed envelope.

The *Phi Delta Kappan* publishes articles concerned with educational research, service, and leadership; issues, trends, and policy are emphasized. Views expressed do not necessarily agree with positions taken by Phi Delta Kappa, the professional fraternity in education.

The Nation's Report Card Goes Home: Good News and Bad About Trends in Achievement

When data from different sources conflict, the truth is often to be found somewhere in the middle, Messrs. Linn and Dunbar point out. Some such truths are evident from the surveys of achievement trends discussed here.

BY ROBERT L. LINN AND STEPHEN B. DUNBAR

TEST RESULTS have long been used to judge educational achievement in the U.S. Public interest in test scores appears to be highest during periods when the public is actively seeking education reforms.[1] Not only are test results used as indicators that reform is needed, but they have also come to be viewed as a major mechanism of reform. Thus it should not be surprising that current interest in test scores is at an all-time high.

Each fall headlines report whether average scores on the American College Testing (ACT) Program's college entrance test and on the College Board's Scholastic Aptitude Test (SAT) have gone up or down for the nation, the state, and the local district. The press also calls attention to differences among the scores of blacks, Hispanics, and whites and to any change in the size of these gaps. The periodic release of results from the National Assessment of Educational Progress (NAEP), now known as the "Nation's Report Card," attracts similar attention. These national reports are supplemented by reports of the results of state- and district-level testing programs and by occasional reports of international assessments.

All in all, there appears to be an abundance of data from which to judge educational achievement. However, the interpretation of the results is quite variable, sometimes apparently contradictory, and frequently subject to controversy. Within the past year or two citizens of most states were apt to read reports indicating that students in their state or district continued to score above the national average and had scored better than their counterparts in the previous year. Those same citizens were also likely to come across reports indicating that scores on the ACT and on the SAT exhibited almost no change, an outcome that U.S. Secretary of Education Lauro Cavazos called "disheartening."

From the NAEP, the public would have heard both good news and bad. The details of the NAEP reports vary some-

"At this rate, I'll never be President, and you'll never be Millie."

ROBERT L. LINN is co-director of the Center for Research on Evaluation, Standards, and Student Testing and a professor of education at the University of Colorado, Boulder. STEPHEN B. DUNBAR is an associate professor of measurement and statistics at the University of Iowa, Iowa City. *The authors wish to thank H. D. Hoover for valuable insights and data pertaining to this article.*

what by the content area of the assessment and by the age of the students, but, by way of overall summary, substantial gains in achievement were recorded by black and Hispanic students during the 1980s, while the achievement of white students changed relatively little. The absolute level of achievement, especially when put in the context of an international comparison released in January 1989, leaves much to be desired.

Some of the apparent discrepancies in messages being sent by test results can be reconciled by taking into account differences in who takes the tests (e.g., self-selected college-bound students or a cross section of high school or elementary school students), in the content of the tests, and in the conditions under which the tests are administered. For example, a substantial part of the highly publicized decline in SAT scores that occurred during the late 1960s and throughout the 1970s can be explained by an expansion of the pool of high school students who were applying to college and taking the test during this period. This is evident from the display in Figure 1 of means for the SAT and for the Preliminary Scholastic Aptitude Test (PSAT).[2] The PSAT results are based on the national norming studies that were conducted in 1960, 1966, 1974, and 1983 and thus are results for comparably defined groups of students. The SAT results, on the other hand, are based on the groups of students who happened to take the test in those four years and hence are not comparable. Because of the changing pool of test-takers, the SAT results clearly give a distorted picture of trends in student performance for the nation as a whole.

Changes in performance in one content area obviously need not match those for another. A gain in math or science achievement, for example, might occur during the same period in which there was a decline in achievement in civics or history. Indeed, it is often suggested that high-stakes testing in limited content areas may lead to an overemphasis on those areas at the expense of content areas not covered by the high-stakes tests.[3]

The average achievement of blacks and of Hispanic students is substantially higher now than a decade ago.

More subtle – and more difficult to explain to the general public – are changes within a content area that affect the meaning and importance of performance on specific test items. As any teacher knows, changes in what is taught greatly affect the specifics of what students know and can do. For example, at the turn of the century it was not unusual for tests required for graduation from the eighth grade to emphasize arithmetic calculation (including such problems as the hand calculation of square roots) and the memorization of names, dates, and places in history. Questions that assess such knowledge and skills would have little value for judging the comparative success of students who have studied a history curriculum that is issues-oriented and a mathematics curriculum that emphasizes conceptual understanding and problem solving and that makes heavy use of calculators.

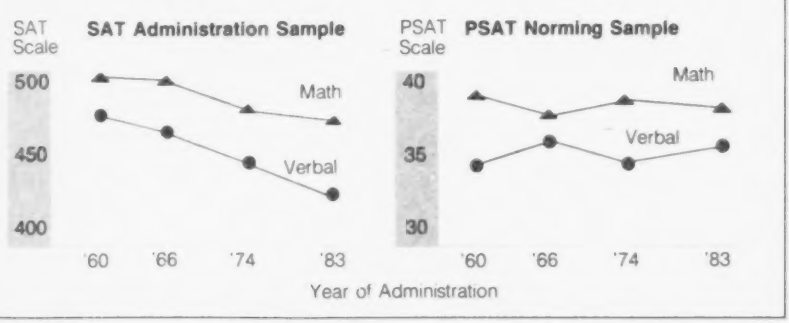
Differences in the purposes of testing and the conditions of testing can also influence results. The NAEP does not report scores for individual students, schools, or even school districts. Indeed, obtaining state-level results from the NAEP became possible only recently.[4]

These circumstances are quite different from the circumstances surrounding typical administrations of achievement tests by states or districts. In such testing programs, not only are individual student scores reported, but, in the case of high-stakes testing, schools and individual teachers are also held accountable for the results. Students' motivation to do well when the scores are sent home and teachers' motivation to teach to the content of the test when they are being evaluated in terms of their students' scores are quite different from the conditions that apply during an administration of the NAEP.

Differences in test security can also partly explain why state and district reports of test results have generally painted a much more positive picture of progress in educational achievement than is provided by the NAEP. There is good reason to believe that the more positive picture derives, in part, from such factors as the exclusion of some students from testing and students' familiarity with the specific content of the tests (since the same forms of the tests are used year after year).[5] There is also reason to believe that the testing conditions in the administration of the low-stakes NAEP may result in an overpessimistic view of student achievement.

After taking into account these differences in the populations of test-takers, in the content of the tests, and in the conditions of test administration, there still remains a mixture of good news and bad. As we will show below, there has been some progress in achievement, most clearly at the low end of the scale. Progress at the high end of the scale is more difficult to ascertain because the results are equivocal. Clearly, the good news is that the average achievement of blacks and of Hispanic students is substantially higher now than it was a decade ago. But

FIGURE 1.
Average SAT/PSAT Performance by Year



the gap between the achievement of minority students and that of white students remains large.

NAEP RESULTS

Perhaps the best single source of information about trends in educational achievement is the National Assessment of Educational Progress. The NAEP avoids the problem of changing pools of test-takers, which plagues the SAT and the ACT, by assessing nationally representative samples of in-school 9-, 13-, and 17-year-olds. While the NAEP content is not designed to match any particular curriculum, it is based on a consensus that attempts to insure that it reflects "the thinking of a wide variety of individuals – state and district curriculum specialists, teachers, school administrators, researchers, parents, concerned citizens, officials, and business people."[6] The NAEP is administered under uniform conditions that protect the security of the tests. Finally, the NAEP includes assessments in several key content areas, including reading, writing, mathematics, science, civics, history, computer literacy, and geography.

Reading. Because of the central role of reading in all school subjects and in the development of an educated citizenry, reading is the most frequently assessed content area in the NAEP. Results are now available from the five assessments conducted between 1971 and 1988. Although the average 9-year-old and the average 17-year-old in 1988 had significantly higher reading performance than their counterparts in 1971, the overall increases were relatively small, and the performance of 13-year-olds has remained essentially unchanged since 1971. Furthermore, the gains that were achieved by 9-year-olds occurred between 1971 and 1980, and the average achievement levels in 1984 and 1988 were slightly lower than the level obtained in 1980.[7]

To go beyond this relatively sterile global summary of the results, it is necessary to see how the best and poorest readers in 1988 compare to their counterparts in earlier years. It is also essential to get a sense of what the scores mean in terms of what students can and cannot do. Finally, it is important to consider separately the performance of minority students, who have traditionally been poorly served by schooling and consequently have recorded lower average levels of achievement. We will go back only to 1975 for this discussion, because the 1971 assessment used a different way of breaking down the sample into racial and ethnic groups than was used in subsequent assessments.

> **The good news is that by 1988 the percentage of black 17-year-olds scoring at 300 or above had increased from 8% to 26%.**

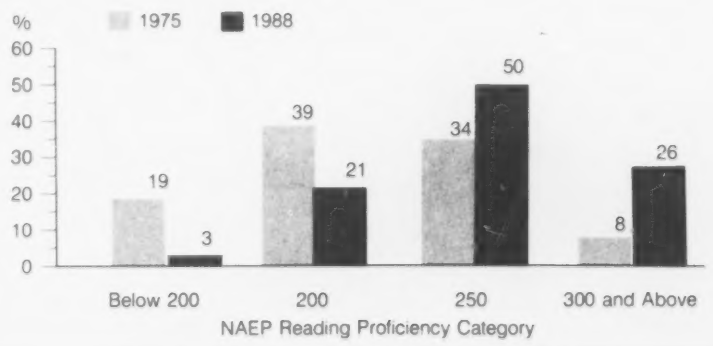
The NAEP reading scale is anchored with example questions and descriptions at five levels. We will focus on the middle three of these: scores of 200, 250, and 300. The NAEP refers to these levels as basic, intermediate, and adept, but such labels do not convey an overall sense of the reading proficiency of students scoring at these levels.

According to the standard NAEP descriptions,[8] students who score at the 200 level "have learned basic comprehension skills and strategies and so can locate and identify facts from simple informational paragraphs, stories, and news articles." They can also "combine ideas and make inferences from short, uncomplicated passages." Students who score at the 250 level "can search for, locate, and organize the information they find in relatively lengthy passages and can recognize paraphrases of what they have read. They can also make inferences and generalizations about main ideas and about an author's purpose." Students who score at the 300 level "can understand complicated literary and informational passages, including material about topics they study at school. They can also analyze and integrate less familiar material and provide reactions to and explanations of the text as a whole."

The anchor points of the NAEP scale are chosen in such a way that a student who scores at the 250 level is highly likely to answer correctly a question at the 250 level or below and is unlikely to answer correctly a question at the 300 level. So when it is said that 50% of the nation's 13-year-olds performed at the 250 level, it doesn't follow that none of the remaining 50% could answer any of the 250-level questions – only that they were unlikely to answer correctly any given question at that level.

It is startling to realize that 19% of 17-year-old black students had levels of reading achievement below 200 in 1975 and that only 8% of 17-year-old black students scored at the 300 level or above. The good news (presented in Figure 2) is that by 1988 the percentage of black 17-year-olds scoring below 200 had decreased from 19% to 3%, while the percentage scoring at 300 or above had increased from 8% to 26%.[9] These results, together with comparable im-

FIGURE 2.
Distribution of Black 17-Year-Olds In Four Reading-Proficiency Categories



provements for Hispanic 17-year-olds between 1975 and 1988, show that the whole distribution of reading proficiency scores of black and Hispanic 17-year-olds shifted to higher levels. That is, not only were average scores higher, but the levels achieved by the top and bottom readers within both of these minority groups increased in absolute terms and in comparison to the performance of their white peers.

White students, on the other hand, made only small gains during this period. But white students constituted the great majority of students in the two highest categories (86% at or above 250 in 1975, rising to 89% in 1988). Since the comparable figures for black and Hispanic 17-year-olds in 1988 were 76% and 73% respectively, it is also clear that there is still a substantial gap between the reading proficiency of the average black or Hispanic 17-year-old and that of the average white 17-year-old. Insuring that all students progress and that the achievement gap between minority and majority students continues to close must be a national priority in the years to come.

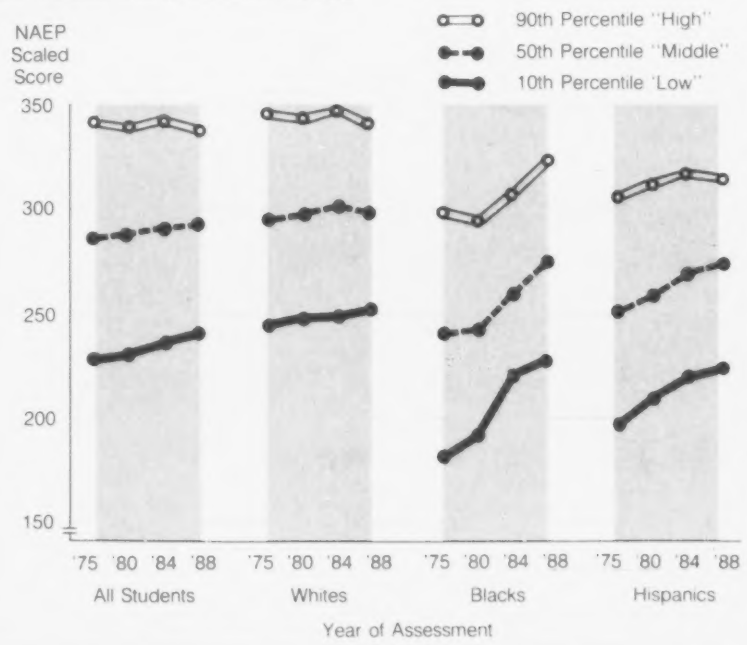
Figure 3 provides a comprehensive picture of the progress of 17-year-olds over the last four assessments in terms that are more familiar outside the context of the NAEP and that can be more readily compared to other data sources.[10] The scores achieved by high-, middle-, and low-scoring students within each of the three major racial/ethnic groups and for the total group are plotted for each of the last four assessments. *High* is here defined as the 90th percentile for a given group and year: that is, the score below which 90% of the students in a given group are found. *Low* is defined as the 10th percentile for the group, and *middle* is defined as the group median.

There is a marked contrast between the trend lines for minorities and those for whites at all three levels. Black students – and to a lesser extent Hispanic students – have demonstrated unequivocal gains in reading comprehension, especially since the 1980 assessment. White students, on the other hand, have an achievement pattern much like that of the total sample: a slight downward trend among high-achieving students, coupled with only modest increases among middle- and low-achieving students. These and similar results at the other two age levels and in other content areas have led to statements in the NAEP reports that the gains have generally been at the more basic levels rather than on tasks requiring higher-order thinking. As we discuss below, the continued narrowing of the gap between minority and nonminority students – rather than the overall trends in the total sample – is bolstered by evidence from other sources of achievement data for the nation.

The trends in reading achievement for 9- and 13-year-olds are more mixed than those for 17-year-olds. Between 1975 and 1980, black, Hispanic, and white 9-year-olds at all achievement levels recorded gains. The performance achieved by the highest-scoring students within each of these groups was also higher in 1988 than in 1980. However, the performance achieved by average- and low-scoring 9-year-olds in 1988 remained relatively constant or in some cases (e.g., low-scoring white students and average-scoring black students and white students) declined slightly.

At age 13, the performance of white students remained essentially unchanged across the four assessments, while the performance of black and Hispanic students improved. The improvement for Hispanic students occurred at all achievement levels – but at less dramatic rates than those posted by 17-year-olds.

FIGURE 3.
NAEP Trends in Reading for 17-Year-Olds At Three Performance Levels



Mathematics and science. Before considering the specifics of trends in achievement in mathematics and science, it is worthwhile to place the issue in the context of national needs and international comparisons. The mathematical needs of the U.S. – indeed, of any advanced technological society – are eloquently and forcefully laid out in *Everybody Counts*, a recent report from the National Research Council. As that report makes clear, mathematics "is the key to opportunity and careers."[11] The importance of mathematics, science, and technology and the need to restructure U.S. precollege education in these areas are also discussed in *Science for All Americans*, a report from Project 2061 of the American Association for the Advancement of Science.[12]

The nation can no longer afford to turn children away from mathematics and science, and the nation can no longer live with the myth that these subjects require "special ability, which most students do not have."[13] One legacy of the myth is that U.S. citizens make up a shrinking fraction of the nation's graduate school enrollment in science and mathematics.

Seniors ranked 14th in advanced algebra, 12th in elementary function/calculus, and 12th in geometry.

Between 1970 and 1987, for example, there was a 50% decline in the number of doctorates in the mathematical sciences awarded to U.S. citizens; by 1987 the majority of new doctoral degrees in mathematical sciences awarded by U.S. universities went to citizens of foreign countries.[14]

In 1981-82 the Second International Mathematics Study (SIMS) obtained achievement test results and a good deal of information about content coverage for eighth-graders from 20 countries (in the case of Japan and Hong Kong, seventh-graders). Test results were also obtained for students from 15 countries who were in their last year of high school. Of the 20 countries participating at the eighth-grade level, U.S. students ranked 10th in arithmetic, 12th in algebra, 16th in geometry, and 18th in measurement. Of the 15 countries participating in the assessment in the senior year of high school, the U.S. ranked 14th in advanced algebra, 12th in elementary function/calculus, and 12th in geometry. Even after taking into account differences in the percentage of students retained through the last year of high school, the level of achievement of U.S. 12th-graders compared unfavorably to that of students in such countries as Japan, Hong Kong, Sweden, Israel, Finland, New Zealand, and England and was similar to that achieved by students in Thailand and Hungary.[15]

Another international assessment involving fewer countries was conducted in mathematics and science in 1988. Samples of 13-year-olds from Ireland, Korea, Spain, the United Kingdom, four Canadian provinces (British Columbia, New Brunswick, Ontario, and Quebec), and the U.S. were assessed using math and science exercises from the 1988 NAEP. Since separate assessments were conducted in French and in English for three of the Canadian provinces, seven separate sets of test results were reported for Canada. Thus 12 populations of students were compared (five countries plus seven province/language groups). U.S. students ranked either last or next to last among these 12 groups on each of five types of problems in mathematics and next to last in each of three categories of science problems.[16]

International rankings are subject to many problems, including sampling, translation from one language to another, differences in retention policies, and differences in curriculum timing. All can cause difficulty in interpreting comparisons such as those just presented. Despite these limitations, the results for the U.S. aroused concern. The 1988 results are particularly sobering in that the exercises were all developed specifically for U.S. students as part of the NAEP. Thus the poor showing of U.S. students cannot be dismissed because the exercises were more in line with the curriculum of another country. In any event, these international results provide a larger context for considering trends in achievement in math and science.

As it does with the reading scale, the NAEP identifies levels of performance in mathematics, anchored at five points on the scale by verbal descriptors and exemplary questions. The three middle levels have been designated beginning skills and understanding (level 200), basic operations and beginning problem solving (level 250), and moderately complex procedures and reasoning (level 300).

> Students who perform at or above level 200 are developing a greater range and depth of basic mathematical skills; however, their use of these skills is still imperfect and relatively inflexible. It can be inferred that learners at this level would have difficulty with reasoning that requires more than simple numerical computation. . . . Students who perform at or above level 250 on the proficiency scale have developed a surface understanding of the four basic operations and are beginning to acquire more complex reasoning skills. . . . Students who perform at or above level 300 demonstrate sophisticated numerical reasoning and are beginning to draw from a wider range of mathematical skill areas, including algebra and geometry.[17]

As in reading, some students at level 200 will answer correctly some questions that exemplify level 250 skills, though they would be unlikely to do so. Conversely, some students at level 300 may answer incorrectly questions that exemplify level 250 skills, though they too would be unlikely to do so.

The only stable trends in the percentage of students at each age level who attain a given proficiency occurred at the 200 and 250 levels. The direction of change at these levels has been positive since the 1978 assessment. While achievement described in this manner has improved, the NAEP notes in *The Mathematics Report Card* (which was based on the 1986 NAEP results) that

only about 21% of the 9-year-olds achieved the level of basic operations and beginning problem solving (level 250), that only about 16% of the 13-year-olds reached the level of moderately complex procedures and reasoning (level 300), and that 6.4% of the 17-year-olds were proficient at skills that were described as "multi-step problem solving and algebra" (level 350). This last finding, coupled with the lack of any upward trend in high-level achievement, prompted much concern in the report of the 1986 assessment over the lack of progress in complex mathematical reasoning.

NAEP data can be used to trace overall trends in mathematics achievement from 1973 through 1986. In general, these data show only minor changes over the 13-year period. For example, the performance of 9- and 13-year-olds rose by a modest (albeit statistically significant) amount. In contrast, the scores of 17-year-olds have shown a net decline since 1973, although improvements since 1982 appear to be stable. As was true for reading, however, the most dramatic changes in the NAEP mathematics findings are the consistent gains made by blacks during the past three assessments. Since 1978 black students have outdistanced their peers in other racial/ethnic groups in the degree of improvement on the NAEP mathematics tests.

The average achievement in the NAEP science assessment increased at all three age levels between 1982 and 1986.[18] But that increase came after a relatively steady decline among 17-year-olds across the four assessments that were conducted between 1969 and 1982, and the subsequent increase for this age group has not been large enough to make up for the earlier decline. The performance of 17-year-olds in 1986 returned to where it had been in 1977 but remained below the level achieved by 17-year-olds in 1969 and in 1973. For 9- and 13-year-olds, the trend line for average achievement, while showing some slight ups and downs between 1970 and 1986, has been nearly flat.

At all three age levels there was a narrowing of the gap between the average science achievement of white students and that of black students – though the relative gains were less dramatic than those posted in reading by black 17-year-olds. The gap in average achievement between Hispanic and white students also narrowed slightly between 1977 and 1986 at ages 9 and 13, but not at age 17. Despite such gains, there remains much to be done at national, state, and local levels to eliminate the still-substantial group differences in achievement for minority students in both science and mathematics.

ADDITIONAL DATA

Besides the NAEP and the college admissions testing programs, data on trends in achievement are available from nationally standardized achievement tests and from the records kept by various statewide testing programs. Perhaps the most complete historical data on educational achievement, analyzed in a recent report from the Congressional Budget Office,[19] come from the Iowa Testing Programs in the form of trend lines for both the nation and the state of Iowa. A renorming of the Iowa Tests of Basic Skills (ITBS) during the 1987-88 school year permits us to examine trends in a nationally representative sample tested under standardized conditions during roughly the period of time covered by the NAEP reading and mathematics assessments.

The trend lines in Figure 4 show a somewhat different picture of achievement in reading and mathematics than was apparent with the NAEP. The performance of a nationally representative sample of third-graders used to standardize the ITBS has shown steady improvement since 1971 and, in contrast with the NAEP, has shown gains at even the highest achievement levels. Although the well-publicized decline in achievement test scores continued well into the Seventies for the seventh-graders, the ITBS data indicate the decline to have ended for seventh-graders by 1978. In reading and mathematics, the rise begun in 1978 continued into the most recent renorming of the ITBS. Indeed, national achievement measured by the ITBS now exceeds the highest point registered before the score decline of the 1960s. As in the NAEP, the more dramatic gains are observed among the lower-achieving students; however, increases at the upper ends of the score distributions are noteworthy in their consistency across all content areas.

The performance of minorities has also shown marked improvement since the 1977-78 standardization of the ITBS. By the time of the 1984-85 standardization, for example, black students in grade 5 had made up about 30% of the gap between their achievement and that of their white peers in the verbal skills measured by the ITBS (vocabulary, reading, language usage, spelling, and so on). However, no consistent narrowing of the achievement gap was registered in quantitative skills (math concepts, problem solving, and computation).

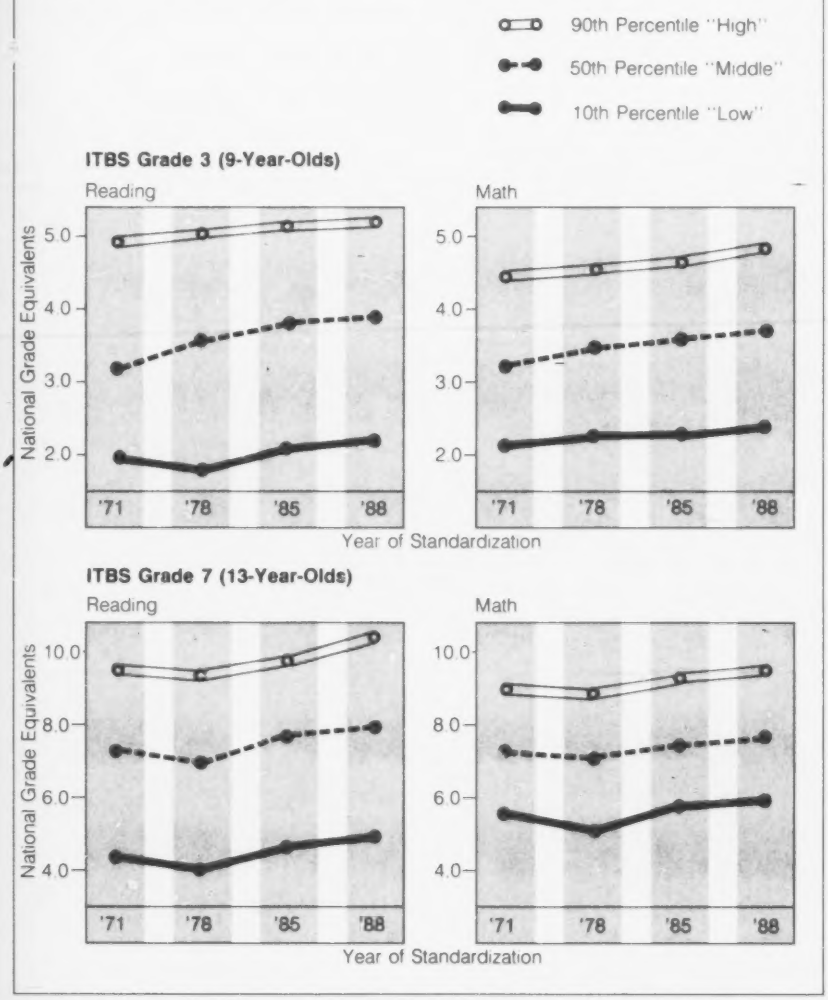
A recent University of Iowa study sheds additional light on the NAEP findings that progress in higher-order skills has been limited. The Iowa study examined the performance of comparable samples of 9-, 13-, and 17-year-olds on test questions in a given subject that were classified as measuring either higher- or lower-order skills.[20] Between 1979 and 1985, performance increases were found on both types of questions. However, contrary to the interpretations of the NAEP, the increases were slightly larger for the complex, higher-order skills. The authors of this study argued that there was little empirical evidence that increases in achievement during the period were limited to relatively mechanical tasks.

WHEN DATA from different sources conflict, the truth is often to be found somewhere in the middle. Some such truths are evident from the surveys of achievement trends discussed above. The gains among low achievers and minorities, in particular, are clear and consistent for all sources of data. What trend data say about the performance of students in, say, the top 10% of their classes is less clear, as are the correct inferences for policy making that can be drawn from any of these apparent trends. Other writers have urged that the education community not presume that a given reform effort will have a desired effect on achievement when the causes of achievement trends are multifarious.[21]

We urge similar caution. But we also suggest that the most recent round of report cards sends several messages to the public.

Reports of achievement trends can be very much influenced by the particular definition of achievement that a test represents and by the conditions under which the test is administered. These factors no doubt account for some of the discrepan-

FIGURE 4.
ITBS Trends at Three Performance Levels



cies detailed in this article. However, it is through a diversity of data on achievement trends that the most accurate, albeit more complex, portrait of educational achievement can be painted. To maintain this diversity is to broaden our understanding of the complexities of educational development and to increase the likelihood of detecting any unanticipated effects of reforms.

Despite the potential flaws of international surveys of achievement, the perspective they provide in mathematics and science suggests that efforts be made in these areas to stimulate and maintain the interests of young people at all levels of achievement and in all subgroups of the student population. What progress has been made in the sciences and mathematics must be carefully nurtured to see that the gains of the Eighties continue and even accelerate in the Nineties. We also have reason to nurture what have come to be labeled higher-order reasoning and problem-solving skills, in both the verbal and the quantitative domains. It is important that such skills be properly represented in the definition of achievement that is embodied in all forms of assessment, so that progress in this area, too, can be accurately measured.

1. Daniel P. Resnick, "History of Educational Testing," in Alexandra K. Wigdor and Wendell R. Garner, eds., *Ability Testing: Uses, Consequences, and Controversies, Part 2* (Washington, D.C.: National Academy Press, 1982), pp. 173-94.

2. *ETS Policy Notes: News from the ETS Policy Information Center*, vol. 2, October 1989, p. 6; and Thomas F. Donlon, ed., *The College Board Technical Handbook for the Scholastic Aptitude Test and Achievement Tests* (New York: College Entrance Examination Board, 1984), p. 176.

3. Edward Haertel, "Student Achievement Tests as Tools of Educational Policy: Practices and Consequences," in Bernard R. Gifford, ed., *Test Policy and Test Performance: Education, Language, and Culture* (Boston: Kluwer Academic Publishers, 1989), pp. 25-50.

4. P.L. 100-297, Part C, provides for the collection and reporting of state-level NAEP results on a trial basis in 1990 and 1992, with voluntary participation.

5. Robert L. Linn, M. Elizabeth Graue, and Nancy M. Sanders, "Comparing State and District Test Results to National Norms: The Validity of Claims That 'Everyone Is Above Average,' " *Educational Measurement: Issues and Practice*, in press.

6. *Reading Objectives: 1986 and 1988 Assessments, National Assessment of Educational Progress* (Princeton, N.J.: Educational Testing Service, 1987), p. 5.

7. Ina V. S. Mullis and Lynn B. Jenkins, *The Reading Report Card, 1971-88: Trends from the Nation's Report Card* (Princeton, N.J.: National Assessment of Educational Progress and Educational Testing Service, 1990).

8. Ibid., pp. 26-30.

9. Ibid., pp. 63-64.

10. Ibid, pp. 62-65.

11. *Everybody Counts: A Report to the Nation on the Future of Mathematics Education* (Washington, D.C.: National Academy Press, 1989).

12. *Project 2061: Science for All Americans* (Washington, D.C.: American Association for the Advancement of Science, 1989).

13. *Everybody Counts . . .* , p. 10.

14. Ibid., pp. 24-25.

15. Curtis C. McKnight et al., *The Underachieving Curriculum: Assessing U.S. School Mathematics from an International Perspective* (Champaign, Ill.: Stipes Publishing Co., 1987); and M. David Miller and Robert L. Linn, "International Retention and Achievement Rates," *Journal of Research in Mathematics Education*, vol. 20, 1989, pp. 28-40.

16. Archie E. Lapointe, Nancy A. Mead, and Gary W. Phillips, *A World of Differences: An International Assessment of Mathematics and Science* (Princeton, N.J.: Educational Testing Service, 1989).

17. John A. Dossey et al., *The Mathematics Report Card: Are We Measuring Up? Trends and Achievement Based on the 1986 National Assessment* (Princeton, N.J.: Educational Testing Service, 1988), pp. 33, 36, 39.

18. Ina V. S. Mullis and Lynn B. Jenkins, *The Science Report Card: Elements of Risk and Recovery — Trends and Achievement Based on the 1986 National Assessment* (Princeton, N.J.: Educational Testing Service, 1988).

19. Daniel Koretz, *Trends in Educational Achievement* (Washington, D.C.: Congressional Budget Office, 1986). See also *Achievement Trends in Iowa: 1955-1989* (Iowa City: Iowa Basic Skills Testing Program, 1989).

20. Elizabeth A. Witt, Myunghee Y. Han, and H. D. Hoover, "Recent Trends in Achievement Test Scores: Which Students Are Improving and on What Levels of Skill Complexity?," paper presented at the annual meeting of the National Council on Measurement in Education, Boston, 1990.

21. Daniel Koretz, *Educational Achievement: Explanations and Implications of Recent Trends* (Washington, D.C.: Congressional Budget Office, 1987). K